

Introduction to Molecular Modeling and Computer Simulation

S. Ravichandran, Ph.D.

Advanced **B**iomedical **C**omputing **C**enter

Bldg. 430, NCI-Frederick, Frederick, MD 21702

E-mail: sravi@ncifcrf.gov

Tel: X1991

Web: <http://nciiris.ncifcrf.gov/~ravichas/MM/MM.htm>

ABCC: Scientific Applications Report

Welcome to the ABCC Scientific Applications Home Page.

From this page you can get information about our Scientific applications and databases. We are also setting up many of the applications to be accessible directly from this page. To get more information about any topic, or to run an application, select the category(ies).

Category:

- All
- Desktop Applications
- Alignment
- Bioinformatics
- Gene Regulation
- Linkage Analysis
- Molecular Modeling
- Molecular Visualization
- Phylogenetic Analysis
- Protein Analysis
- Quantum Chemistry
- RNA Analysis
- Sequence Analysis
- XRay/NMR Structure Refinement
- Other

Run Report Reset

National Cancer Institute at Frederick
Problems/Comments: Support, Feedback & Help Request
Site Coordinator: ABCC Site Coordinator

Copyright © 2002, ABCC
Advanced Biomedical Computing Center

ABCC

Advanced Biomedical Computing Center

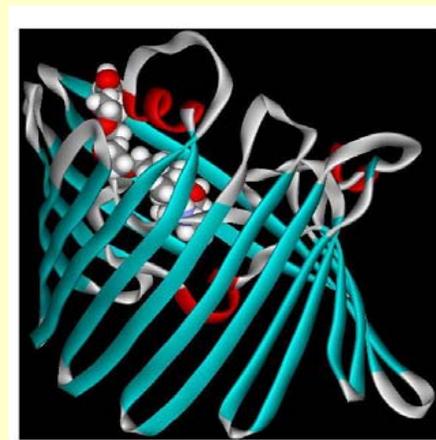
<http://www.abcc.ncifcrf.gov>

ABCC Scientific Applications Page

Software available through ABCC and instructions on how to use them

Overview of the class

- What is Molecular Modeling?
 - Basic assumptions etc.
- Where do we get these models?
 - Model Building, Databases
- Displaying models and model properties?
 - Graphics
 - Properties: Hydrophobicity, electrostatics etc.
- Simulating the models
 - Techniques: MM, MD, QM
- Selected Applications
 - QSAR, Protein-ligand docking, Drug Design
- Hands on exercise: DSViewerPro 6.0
 - Small Molecule building, Energy Minimization
 - Aligning small molecules
 - Protein structure visualization



Porin

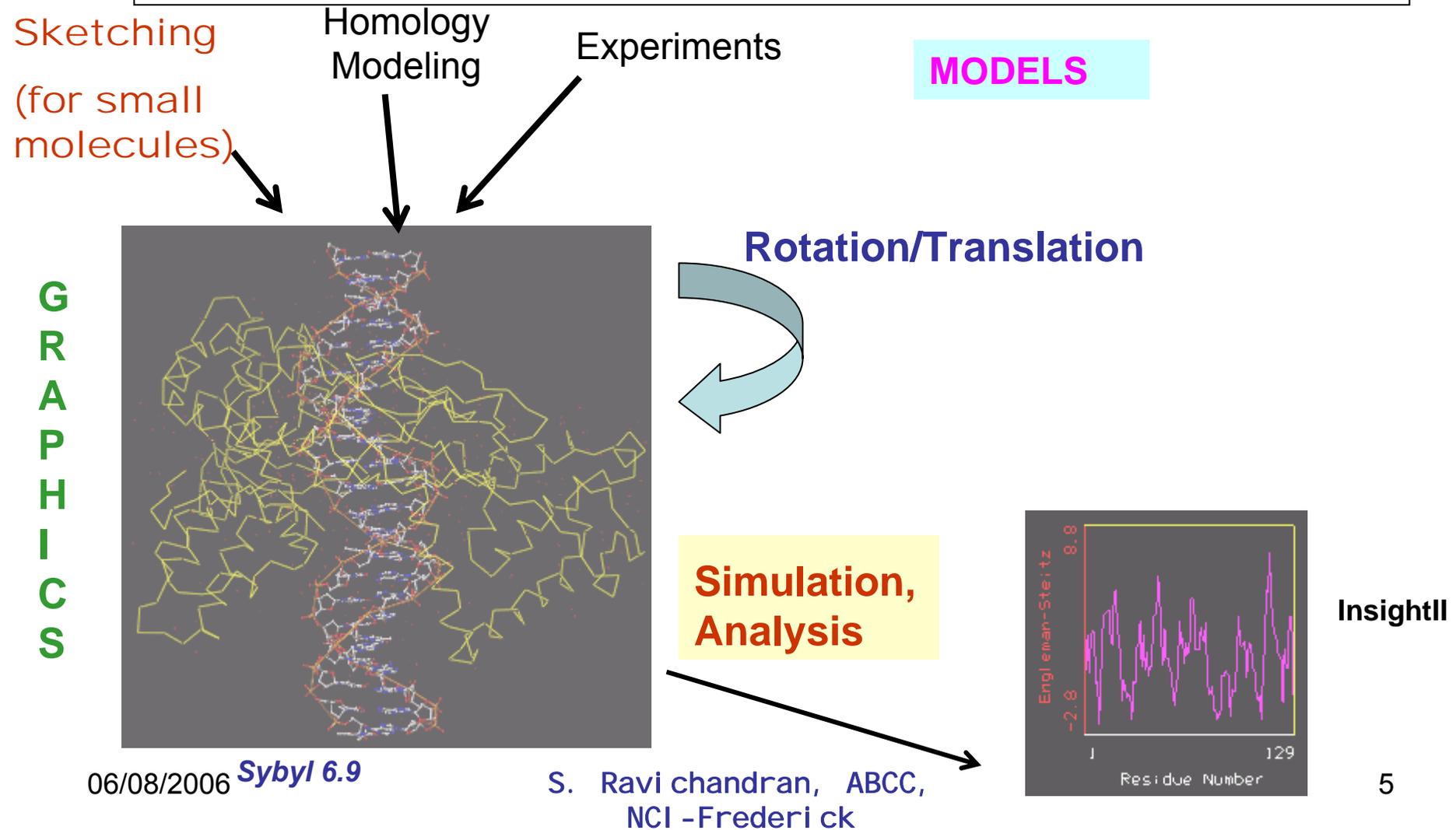
Using DSViewerPro

Models

- Models are central to the understanding of the concepts in Chemistry/Biology
- Models are the result of mathematical equations
 - Classical and Quantum Mechanical
- Models have to be build carefully
 - Experimental information has to be taken into effect during model building

Overview of Molecular Modeling

Definition: Molecular Modeling is the science of creating model structures (numerically) and simulating its function using Classical or Quantum Mechanical laws

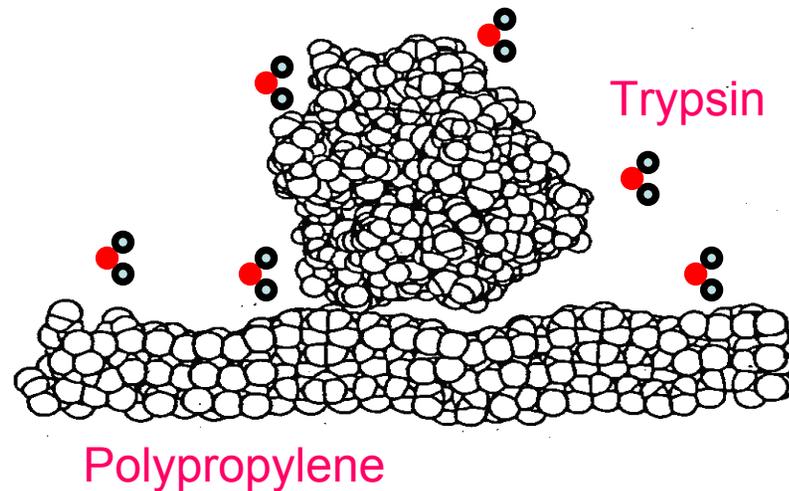


Real Systems

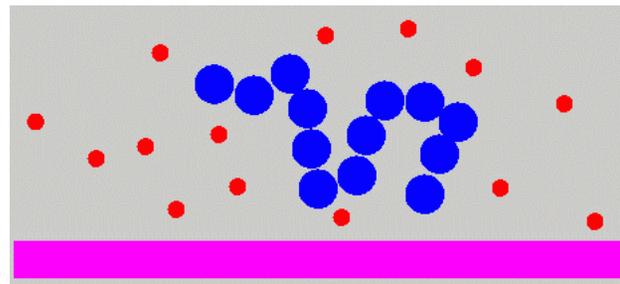
Perform
Experiments

Experimental
Results

Molecular Modeling

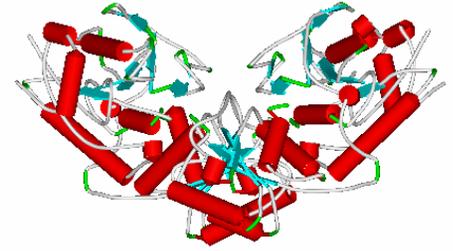


Model to study protein-surface adsorption
 α chain model



Allen & Tildesley

Molecular Modeling: Application



- Computational Biology/Biochemistry/Bioinformatics
 - Milestone: 3D to 1D abstraction
 - Save storage—Imagine the speed of sequencing
 - Represents the nature of a molecule correctly and ignore the details of atomic information



$\Sigma = \{A, C, G, T\}$ DNA

$\Sigma = \{A, C, G, U\}$ RNA

$\Sigma = \{A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y\}$

Alphabets

Molecular Modeling: A simple example



1WSA backbone structure displayed using Sybyl 6.9

Entry 1WSA reports 26 Missing residues-a residues-a flexible loop

Loop is found to be important for the functioning of this molecule.

Advantage: 3-D structure availability

Disadvantage: incomplete structure

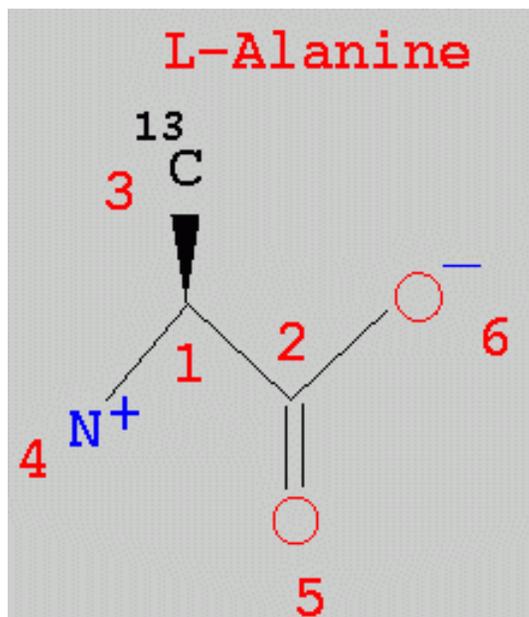
How do we get the complete structure?

Molecular Modeling

Getting Small Molecules

- Databases
 - Cambridge Structural Database (Experimental)
- Sketching
 - User sketches the molecule on the computer and the software converts into a proper 3D molecule (bond-angle, bond length etc.)
 - Sybyl, InsightII.....
- Fragment Libraries
 - Fragments are available in the software library. User builds the molecule using the fragments as a LEGO blocks
 - Sybyl, InsightII,.....
- 2D-3D conversion tools
 - Omega, CORINA, CONCORD etc.
 - ABCC has license for Omega

MDL 2D file format (CTFILE)



Exercise 1: 2D SD File

```

6 5 0 0 1 0          3 v2000          Counts Line
-0.6622  0.5342      0.000 C  0  0  2  0  0  0
 0.6622 -0.3000      0.000 C  0  0  0  0  0  0
-0.7207  2.0817      0.000 C  1  0  0  0  0  0 Atom
-1.8622 -0.3695      0.000 N  0  3  0  0  0  0 Block
 0.6220 -1.8037      0.000 O  0  0  0  0  0  0
 1.9464  0.4244      0.000 O  5  0  0  0  0  0
1 2 1 0 0 0
1 3 1 1 0 0
1 4 1 0 0 0          Bond Block
2 5 2 0 0 0
2 6 1 0 0 0
M CHG 2  4  1  6  -1  Properties Block
M ISO 1  3  13
M END
  
```

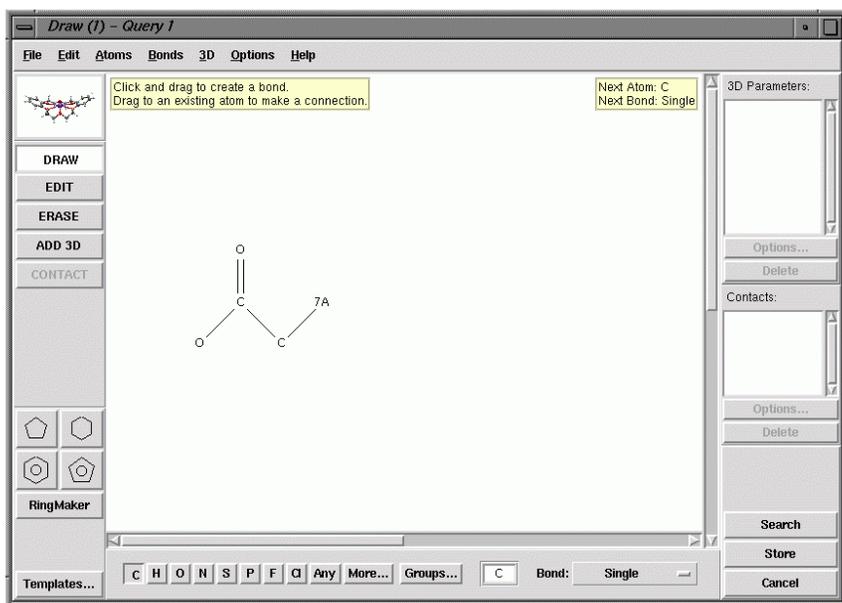
Connection
Table (Ctab)

Cambridge Structural Database (CSD)

- X-ray and neutron diffraction analysis of carbon-containing molecules (up to 1000 atoms including H)
 - Organics, Organometallics, Metal Complexes
 - Peptides up to 24 residues
 - mono-, di- and tri-nucleotides
- Different Search Options:
 - Basic substructure, Substructure with constraints, 3D substructure, non-bonded interactions, Pharmacophore, Cell parameter, Journal Reference

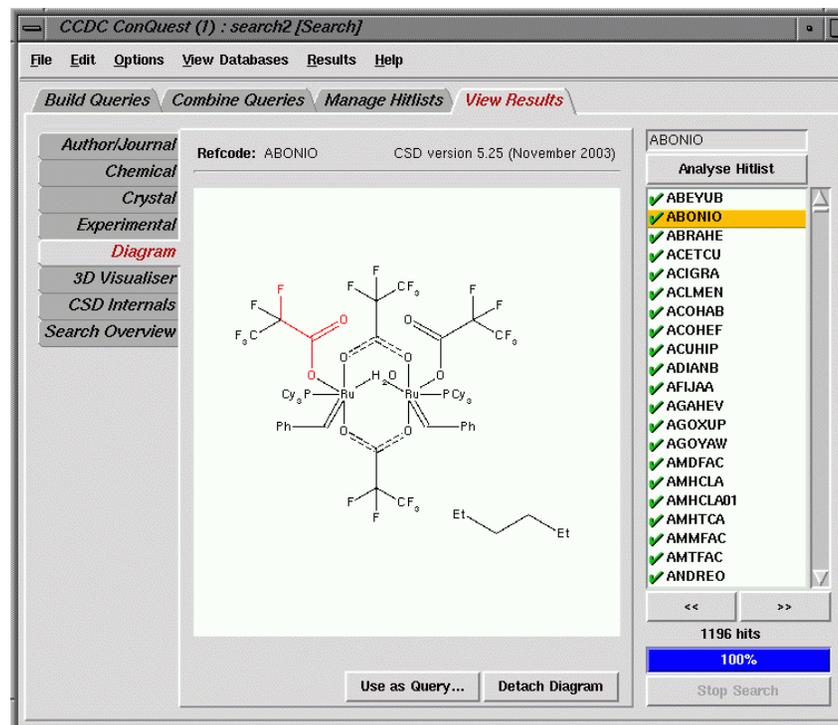
CSD: Substructure Search

Query



7a : any halogen

Result



Pubchem database (NIH)

- Pubchem: Library of small molecules

- Structures, activities

- How to search?

- Compound search

- using names, synonyms, keywords, identifiers

- Substance search for deposited structures

- using names, synonyms, keywords, identifiers

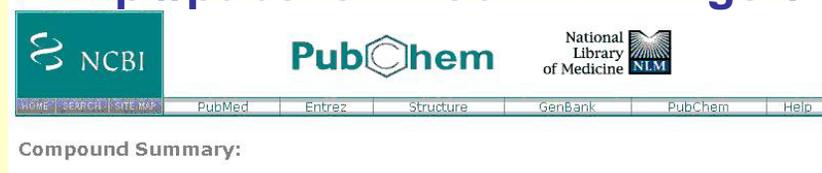
- Bioassay search for description

- using names (ex HIV growth inhibition)

- Structure search

- using example SMILES, mol files

<http://pubchem.ncbi.nlm.nih.gov/>



3D-Structural Database of biomolecules (PDB)

<http://www.rcsb.org>

PDB: 50
structures
(1975)

NMR: 3776
X-ray: 24334
Total: 28110
May 10, 2005

- NMR (early 1960s)
 - Dynamic
 - Multiple Models (Each conformation is a model)
 - Aqueous environment
 - Limitations
 - Size of molecule
 - < 30kD
- Examples
 - [1DV0](#), [1UBA](#)

- X-ray (1958)
 - Static
 - Only one model
 - Crystal
 - Limitations
 - Not limited by size
- Examples
 - [7LYZ](#), [2SRC](#)

NMR: 5409
X-ray: 31223
e⁻ micros: 125
Total: 36837

Jun 01, 2006

Anatomy of PDB file

Atom #

Residue

Residue #

```

ATOM      1  N   MET B   1      52.127 -7.410  40.963  1.00 52.97
ATOM      2  CA  MET B   1      51.096 -6.600  40.340  1.00 52.20
ATOM      3  C   MET B   1      51.305 -6.416  38.831  1.00 51.74
ATOM      4  O   MET B   1      52.405 -6.086  38.379  1.00 50.94
ATOM      5  CB  MET B   1      51.012 -5.245  41.044  1.00 52.27
ATOM      6  CG  MET B   1      50.777 -5.353  42.542  1.00 50.84
ATOM      7  SD  MET B   1      49.355 -6.390  42.934  1.00 51.51
ATOM      8  CE  MET B   1      48.078 -5.169  43.363  1.00 47.75
ATOM      9  OXT MET B   1      50.932 -9.312  40.816  1.00 53.41
ATOM     11  N   ASN B   2      50.235 -6.630  38.064  1.00 50.80
ATOM     12  CA  ASN B   2      50.271 -6.496  36.613  1.00 49.26
ATOM     13  C   ASN B   2      50.332 -5.038  36.246  1.00 48.37
ATOM     14  O   ASN B   2      50.120 -4.673  35.089  1.00 50.24
ATOM     15  CB  ASN B   2      49.016 -7.074  35.977  1.00 49.62
ATOM     16  CG  ASN B   2      48.753 -8.479  36.395  1.00 51.39
ATOM     17  OD1 ASN B   2      49.628 -9.339  36.316  1.00 51.26
ATOM     18  ND2 ASN B   2      47.531 -8.701  36.861  1.00 54.69
.....
HETATM 2462 ZN   ZN   909      45.731   9.445  45.851  0.54 77.21
HETATM 2463 C1  RET B  978      33.234   8.591  25.798  1.00 34.05
HETATM 2464 C2  RET B  978      31.995   8.387  24.968  1.00 33.95
HETATM 2465 C3  RET B  978      32.242   8.645  23.513  1.00 33.45
HETATM 2466 C4  RET B  978      32.720  10.104  23.258  1.00 33.69
HETATM 2467 C5  RET B  978      33.717  10.591  24.302  1.00 34.43
HETATM 2468 C6  RET B  978      33.938   9.900  25.443  1.00 34.95
HETATM 2469 C7  RET B  978      34.915  10.362  26.451  1.00 36.11
    
```

Only portion of the file is shown

Temp Factor

Chain ID

X Y Z Occ

More about PDB Structures

- These crystallographic databases gives information w.r.t a crystal environment
 - Proteins NMR studies have shown that the structure in the crystal phase and solution phase are almost same but for small molecules this may not be the case
 - PDB doesn't cover the whole spectrum because some of the molecules cannot be crystallized

No Experimental Macromolecule Structure & Homology Modeling

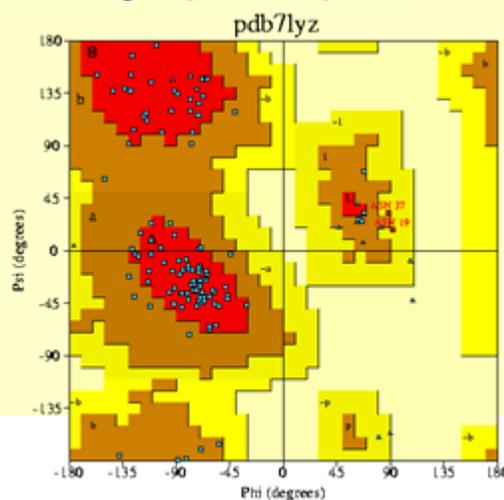
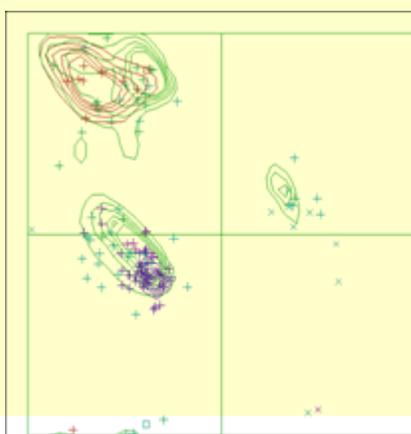
- No 3D structure but has homologous PDB entries
 - Can exploit homology to model the unknown protein
 - Accelrys (Modeller), Swiss-Model, Tripos (Matchmaker)
- No 3D structure but do not have any homologous PDB entries
 - Threading, Reverse Folding
 - Tripos (GenFold)

Quality (model) check!

- **Procheck**: Stereo-chemical quality of the protein and residue by residue analysis in figures

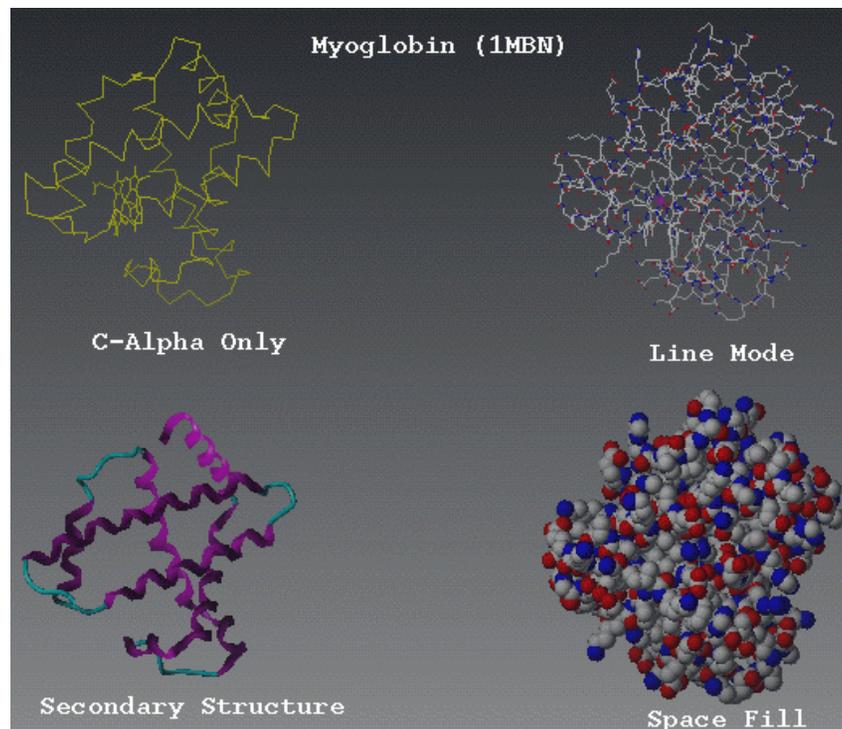
<http://www.biochem.ucl.ac.uk/~roman/procheck/procheck.html>

- **PDBREPORT**: <http://www.cmbi.kun.nl/gv/pdbreport>

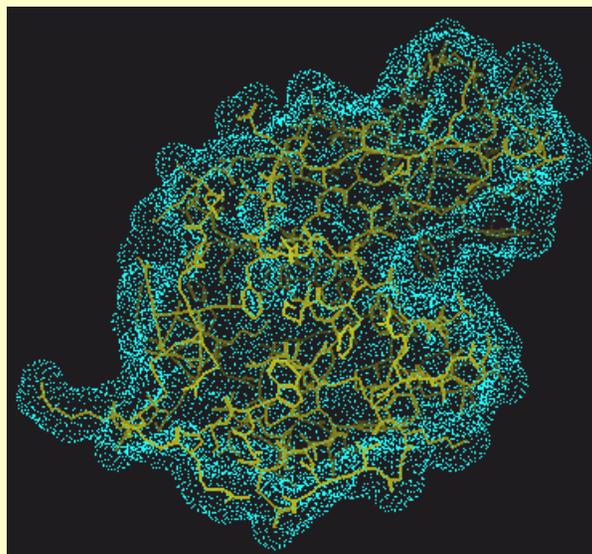


Molecular Modeling: Visualization

- Visualization
 - Free: [Spdbv](#), [Cn3D](#), [Rasmol](#), [VMD](#) and many more
 - \$\$\$\$: [Tripos](#), [Accelrys](#) and many more

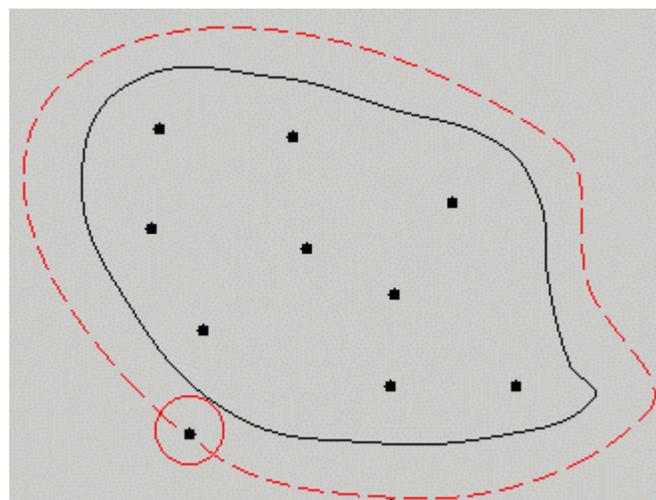
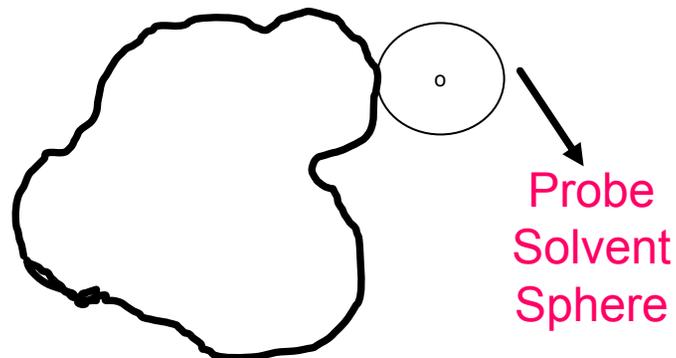
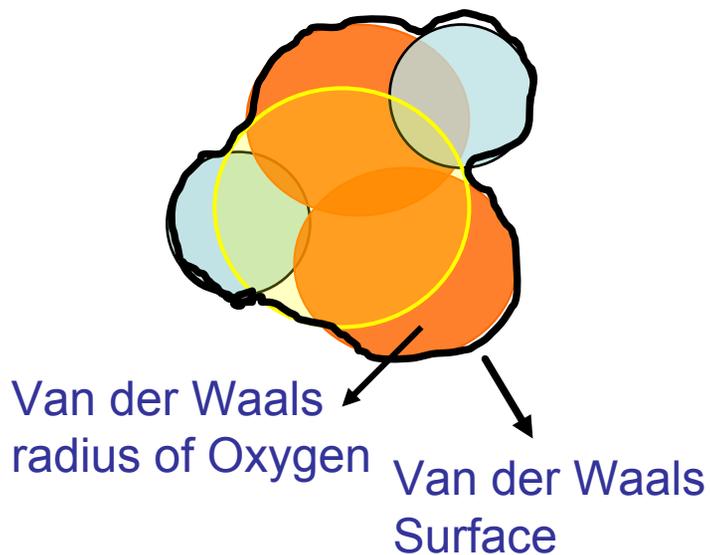
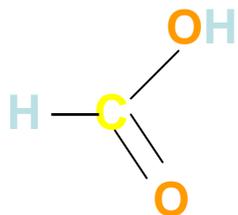


Selected Static properties of Macromolecules: SA

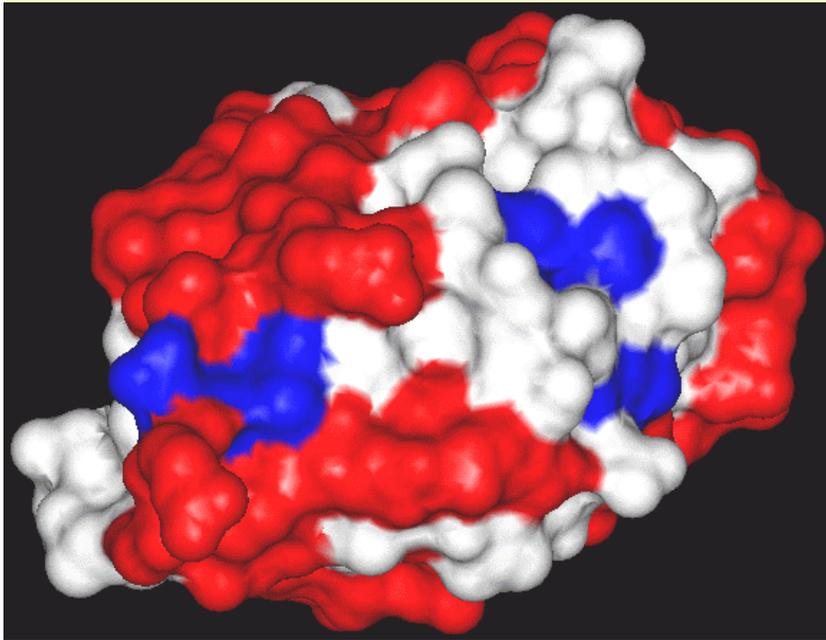


- Solvent-Accessibility (SA)
 - SA help us to know what groups are on the surface-solvent exposed
 - Can give hints on the possible interaction with ligands etc.

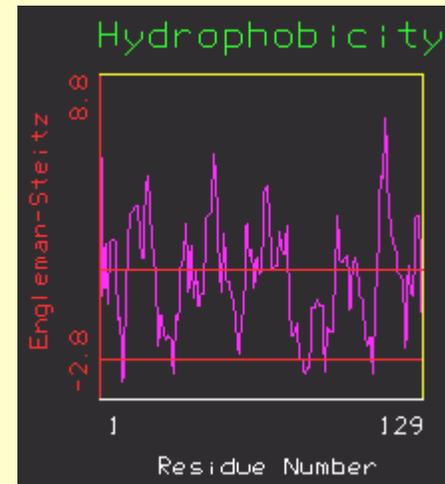
Molecular Surfaces



Molecular Surface



Hydrophobic residues on the Connolly surface of the protein for some reason!
Protein-Protein Interactions



Lysozyme, Hydrophilic red, hydrophobic blue, InsightII, subsets are created using Engleman-Steitz algorithm

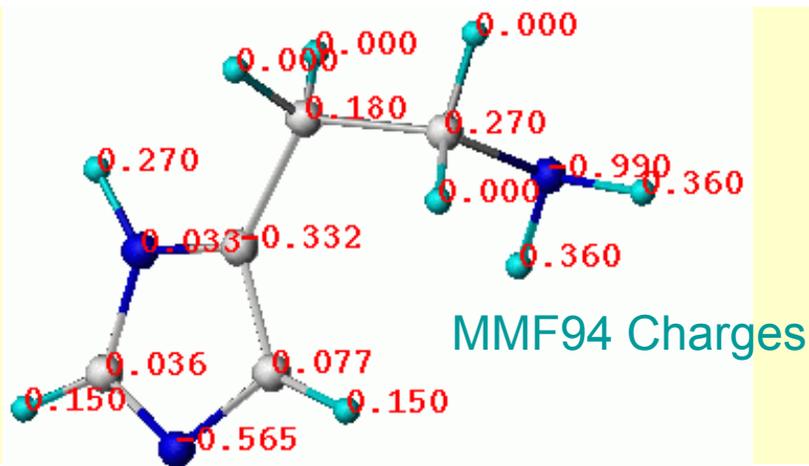
Electrostatics: Point charges/Partial Charges

Topological Procedure (ex Gasteiger-Huckel method)

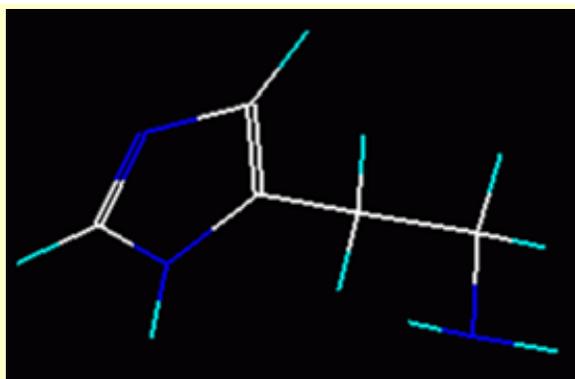
- a) Uses the electronegativity of different atom types.
- b) Do not need structural geometry or conformation of the molecule.
- c) Total charge on an atom is a sum of sigma (σ) and pi (π) component. π (conjugated systems) is calculated first followed by σ component

Disadvantage: Neglect of geometries and conformation. Gasteiger-Huckel do not have atom types for certain atoms (ex. Silicon)

Partial Charges

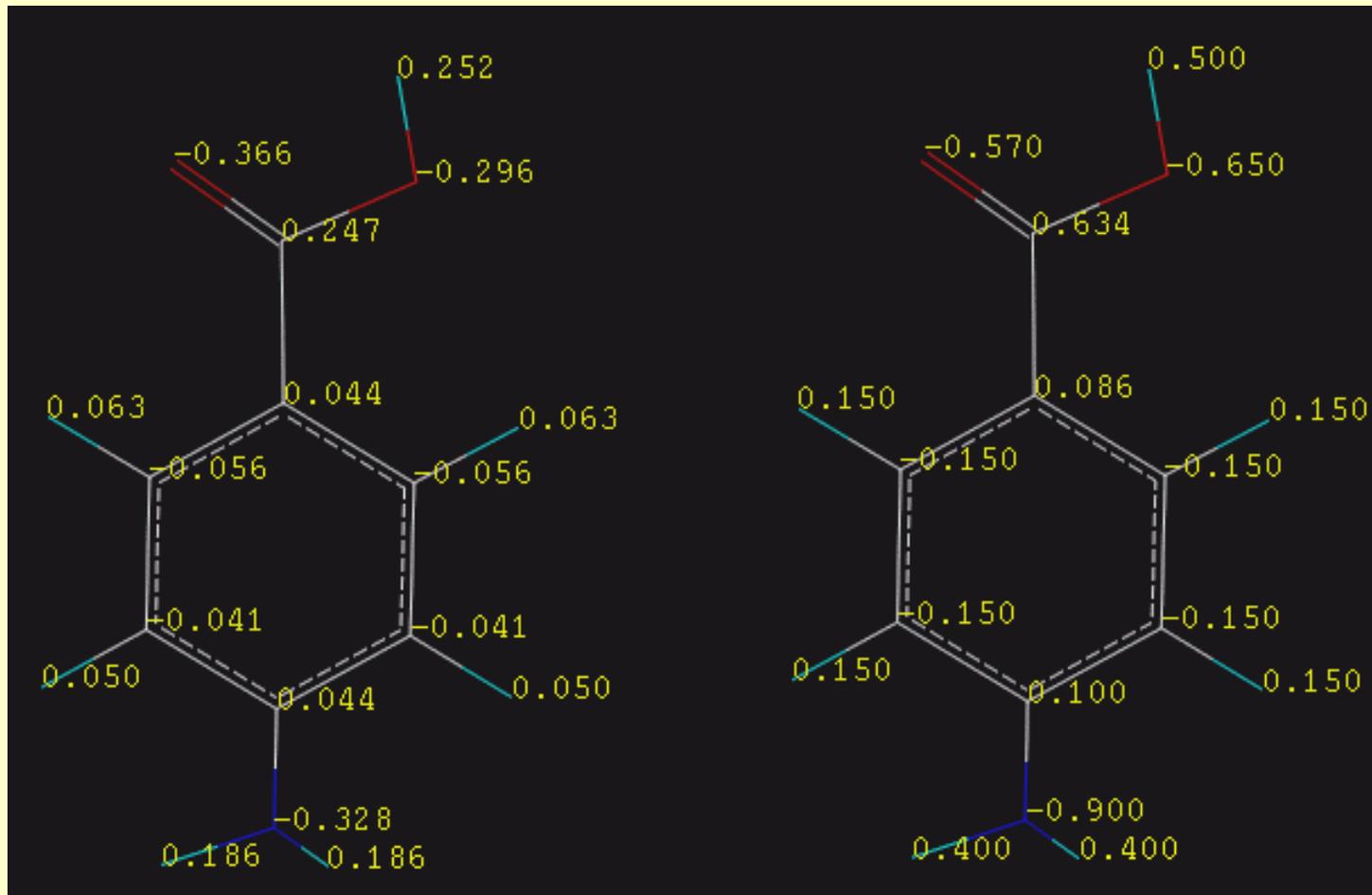


- Classical View
 - Valence electrons fixed to atoms
- Modern View:
 - Diffused
 - Electrons spend more time near electronegative atoms
 - Charges on Nitrogen are different (positive and negative)



Histamine

Partial Charges



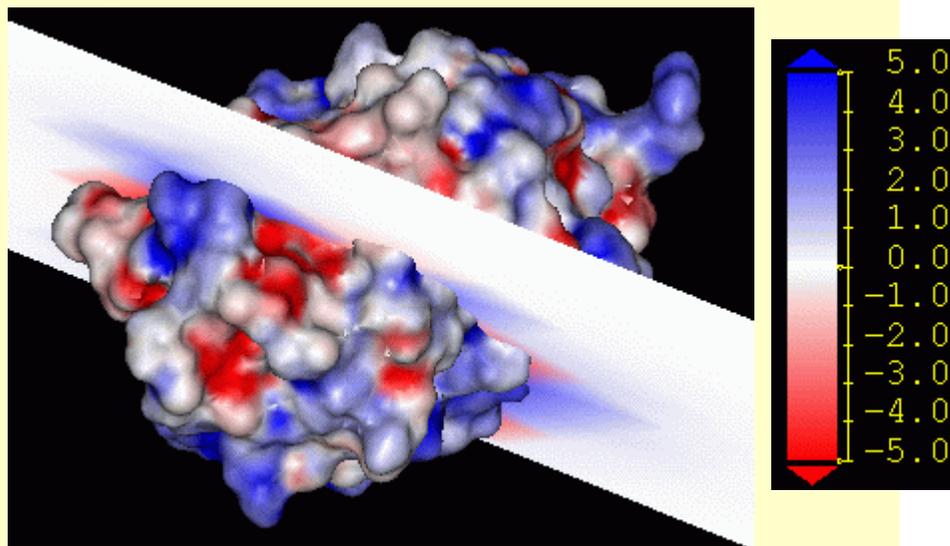
Gasteiger-Huckel

MMF94

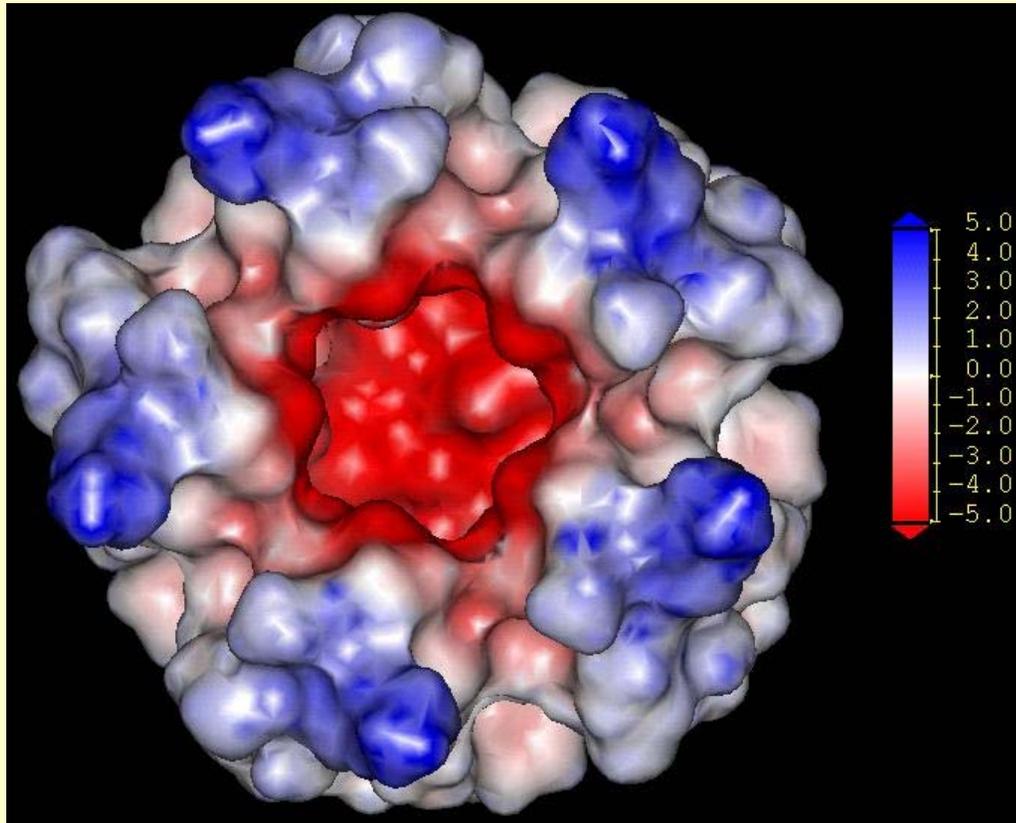
Electrostatics

- Delphi: software to calculate electrostatic properties
 - Protein-Ligand interactions
 - Calculate electrostatic potential
 - Effects of site-directed mutagenesis
 - Electrostatic contribution to the solvation energy

Picture made with InsightII



Electrostatics



Potassium Ion
Channel

Red -ve

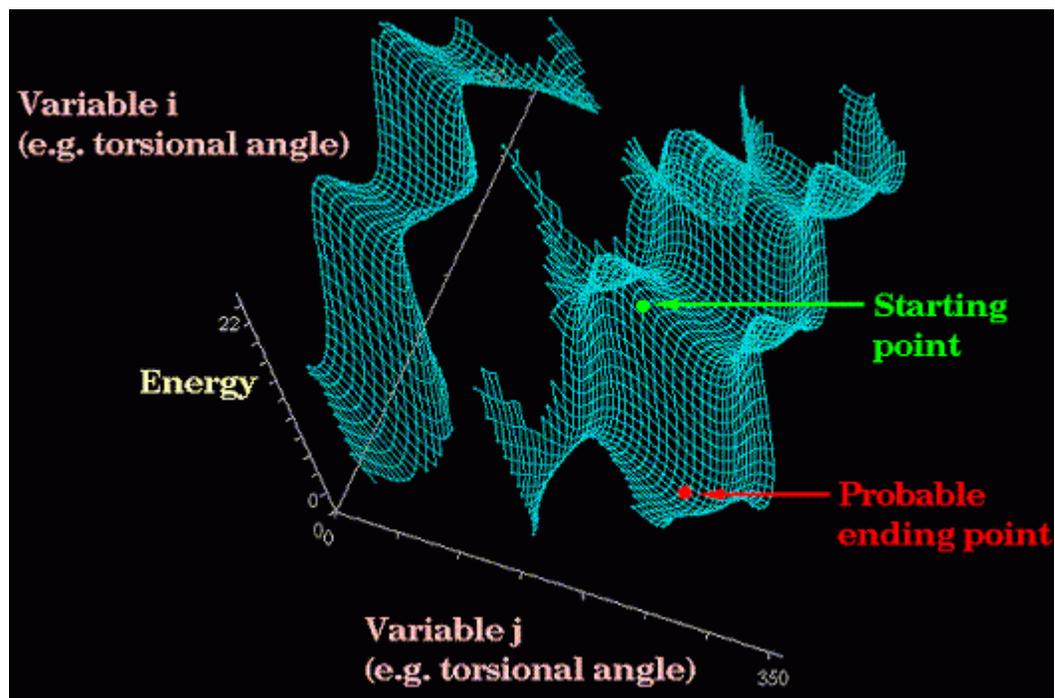
Blue +ve

Computer Simulation

Molecular Mechanics (MM)

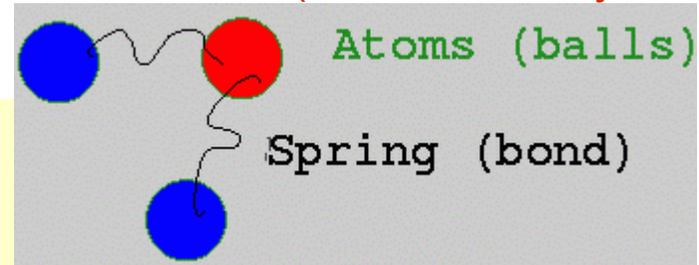
- What is Molecular Mechanics?
 - MM is a energy refinement procedure. Refinement process is usually called Minimization or Energy Minimization.
 - Assumption: Energy minimized structure is closer to the stable geometry and probably closer to experimental structure.
- Where Energy Minimization is usually employed?
 - Molecule Building, Homology modeling, Conformational Search, PDB file refinement

MM



NIH Molecular Modeling

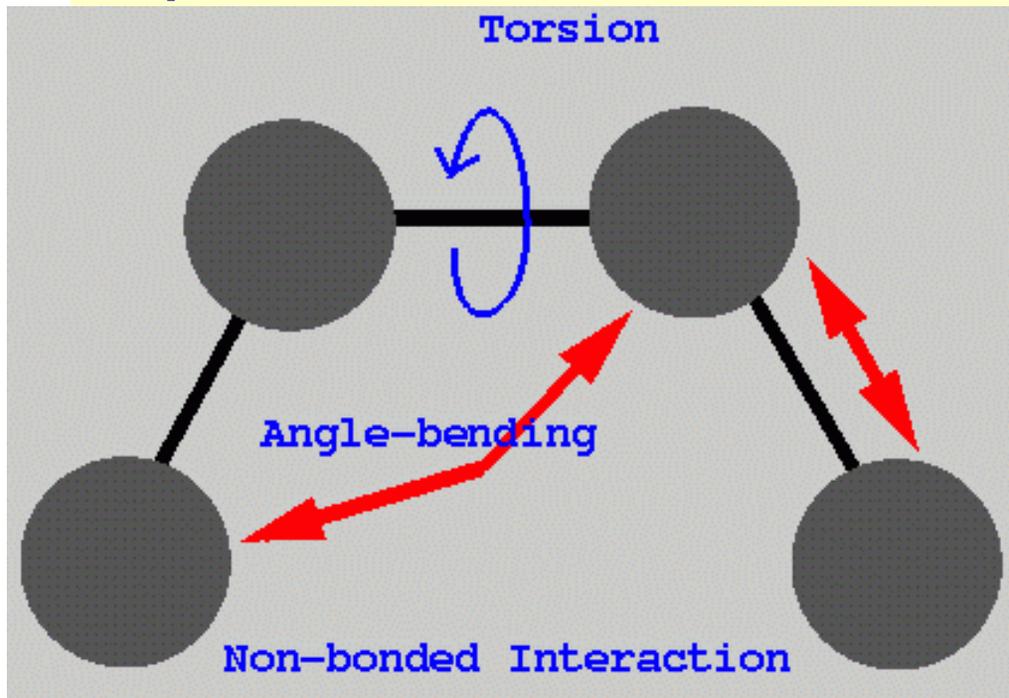
Basic assumptions of MM_(1960's to early 70s)



- Molecules are assumed to be soft balls (point masses) and connected to others by bonds (springs)
- Electrons and nuclei are lumped together
- Total energy of the system is an important property and it is usually computed as a sum of independent energy terms.

ForceField

- ForceField is an analytical Functional form for the independent energy terms and parameters

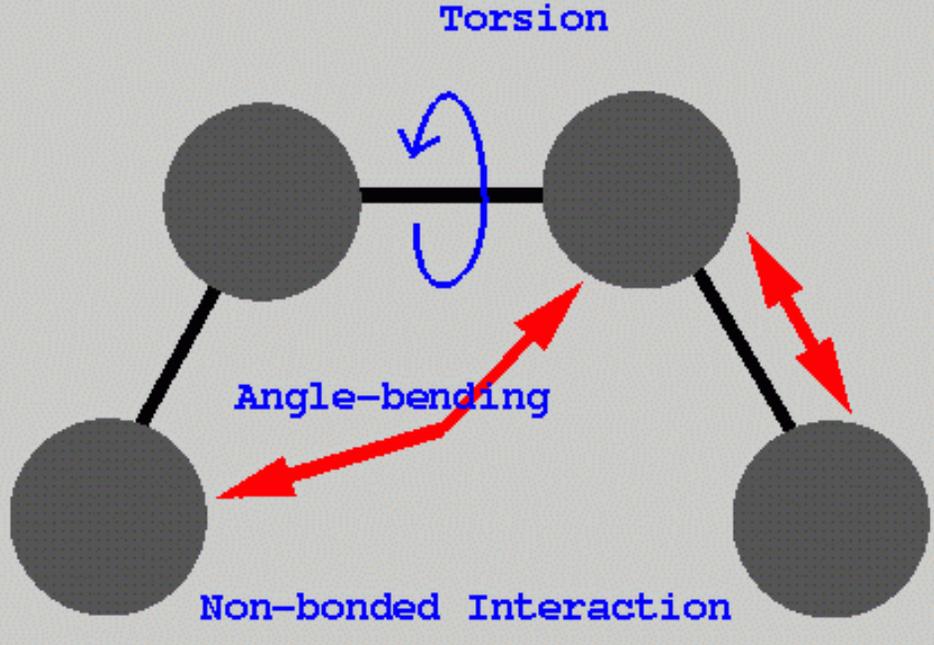


$$E_{\text{pot}} = \underbrace{\sum 1/2 K_b (b - b_0)^2}_{\text{VDW}} + \sum 1/2 K_\theta (\theta - \theta_0)^2 + \sum 1/2 K_\phi (1 + \cos N\phi)^2 + \sum 1/2 K_\chi (\chi - \chi_0)^2 + \underbrace{\sum (B/r)^{12} - (A/r)^6}_{\text{electrostatic}} + \sum (qq/r)$$

Bond Stretching = $K(b-b_0)^2$
Simple Functional Form

Force \rightarrow $-\frac{\partial E}{\partial r_i}$ Hooke's Law

- Additivity
- Transferability of Force Field parameters



ForceField

$$E_{\text{pot}} = \sum 1/2 K_b (b - b_0)^2 + \sum 1/2 K_\theta (\theta - \theta_0)^2 + \sum 1/2 K_\phi (1 + \cos N\phi)^2 + \sum 1/2 K_\chi (\chi - \chi_0)^2 + \sum (B/r)^{12} - (A/r)^6 + \sum (qq/r)$$

Force \rightarrow $\frac{-\partial E}{\partial r_i}$

$$E(\theta) = k_\theta/2 (\theta - \theta_0)^2 \quad (\text{Hooke's Law, Simple form})$$

$$E(\theta) = k_\theta/2 (\theta - \theta_0)^2 [1 - k'(\theta - \theta_0) - k''(\theta - \theta_0)^2]$$

(higher order terms (shown in blue))

Why higher orders?

To increase accuracy, account for larger deformations

MM3, MM4, CFF, MMFF use cubic or quadratic or higher terms.

MM



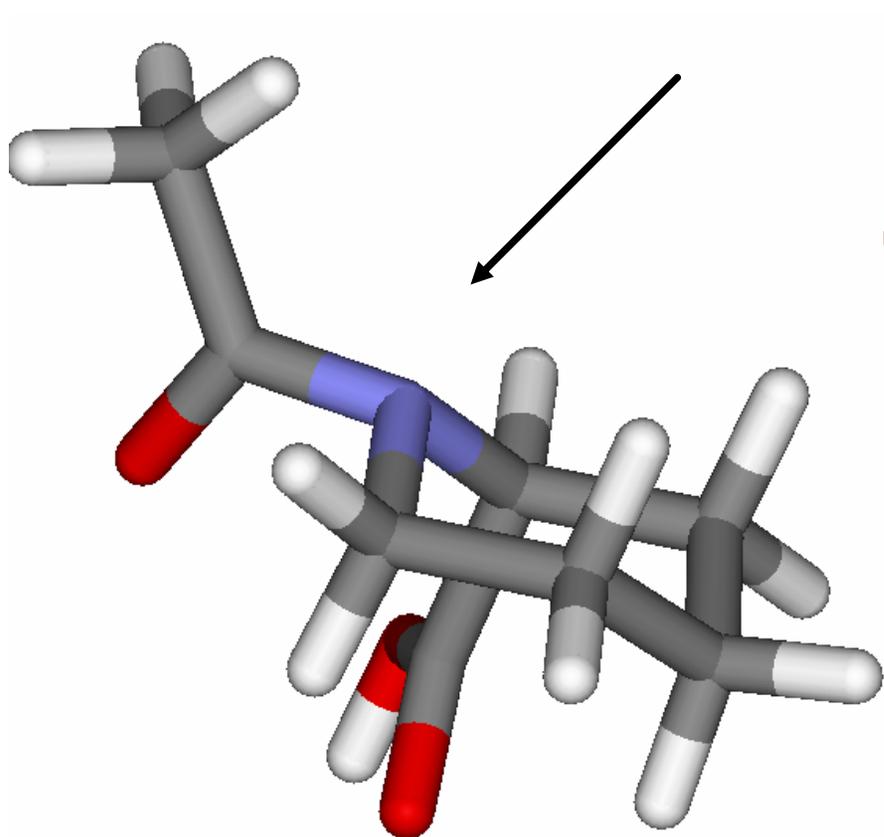
- Each atom/bond in a molecule or amino acid is identified by
 - 1) Atom type 2) Residue type 3) hybridization type 4) Bonding info. 5) Charges 6) Coordinates

Carbon-di-oxide

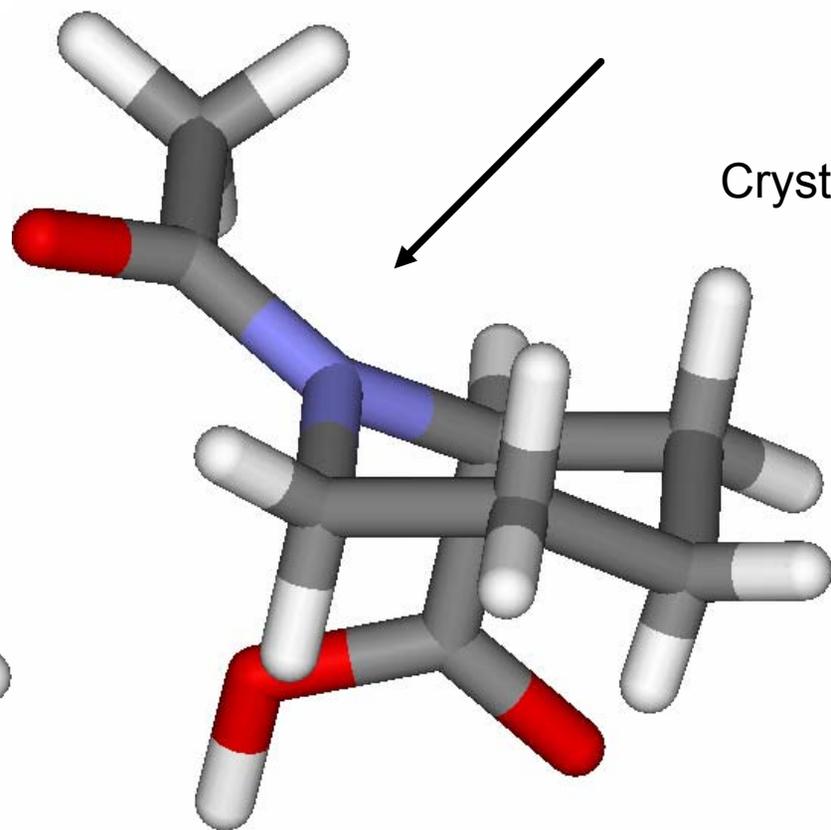
- Atom types C1, O2; Hybridization: sp, sp²; Bonding info: C1 is bonded to two O2 atoms; Coordinates: x,y and z; Charges: (C) 0.372, (O) -0.186

N-Acetyl-Piperidine-2-
carboxylic acid

AtomTypes/FF



N-Sp3



Experiments/QM CalcIns

Crystal Str

Force Field (small molecules)

- Force Field for small molecules
 - Norman Allinger & co-workers (1977)
<http://europa.chem.uga.edu/allinger/mm2mm3.html>
 - MM2 (1991)
 - Parameter update (1991), Functional update (1987)
 - MM3(2000) Latest Version of MM3
 - Extended to handle amides, polypeptides & proteins
 - MM4 (2003)
 - Emphasis on alkanes, non-conjugated alkenes, conjugated hydrocarbons and vibrational frequencies

Force Fields for small molecules

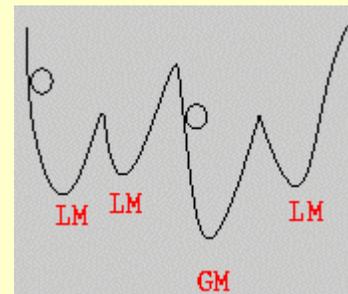
- Other Force Fields for small molecules
 - TINKER
 - Universal Force Field
 - MOMECC
 - Inorganic and coordination compounds
 - COSMOS
 - Computer simulation of Molecular Structures)
 - Bond polarization

Force Field for biomolecules

- **AMBER: Assisted Model Building with Energy Refinement** (Peter Kollman at UCSF & Collaborators with groups from academics and industry)
 - Latest version (AMBER 8)
 - AMBER5 AMBER6: Cornell et al (1994) FF or parm94
- **CHARMm**
 - **Chemistry at HARvard Molecular Mechanics (CHARMm) Martin Karplus (Cambridge MA)**
 - <http://www.accelrys.com>
- **GROMOS (GROningen MOlecular Simulation)**
 - **Wilfred Van Gunsteren & Herman Berendsen**
 - Groningen (Netherlands)
- **OPLS (Optimized Potential for Liquid Simulation)**
- **CVFF/CFF (Consistent Force Field)**
- **MMFF (Merck Molecular Force Field)**
 - **Thomas Halgren, Merck and Co.**

Energy Minimization

- Different Flavours of Energy Minimization:
 - Steepest Descent (SD)
 - SD is used to relieve overlaps and so good at start
 - Conjugate-Gradient (CD)
 - CD is slow but can lead to structures with low energies. Do not get trapped in local minima like SD!
 - Simulated Annealing



MM

– Limitations: No guarantee that you will reach Global minimum

- Two alternative methods:
MD or stepwise rotation of bonds

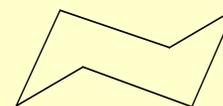
Twist boat
cyclohexane
11.917 kcal/mol

Molecular
Dynamics

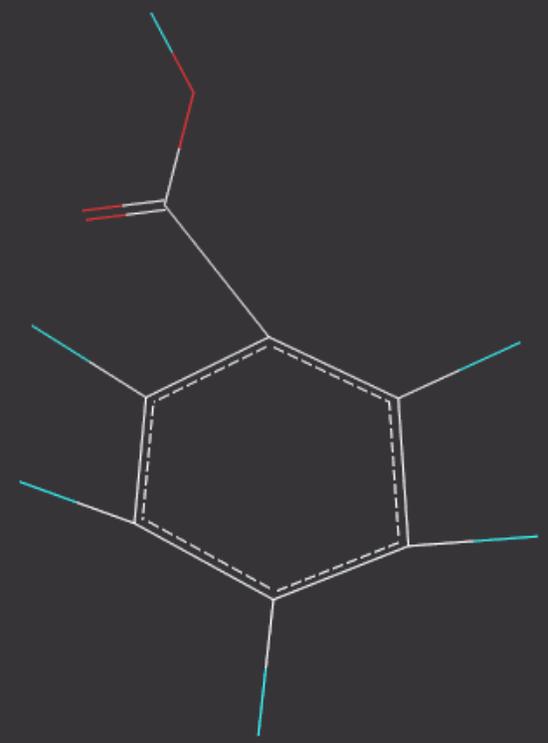
Chair form
6.558 kcal/mol

Example

Cyclohexane
remains in
twist boat form
in Molecular
Mechanics

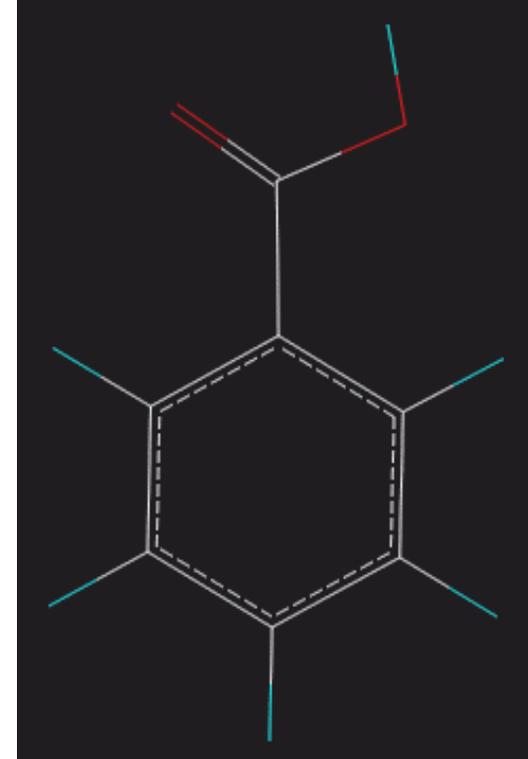


ACD/ChemSketch



Initial
403.252
kcal/mol

Final →
11.453
kcal/mol



Method:
Powell

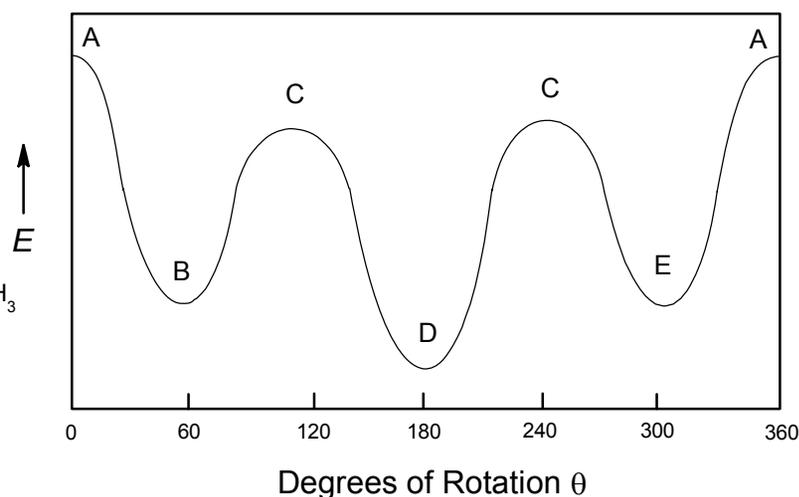
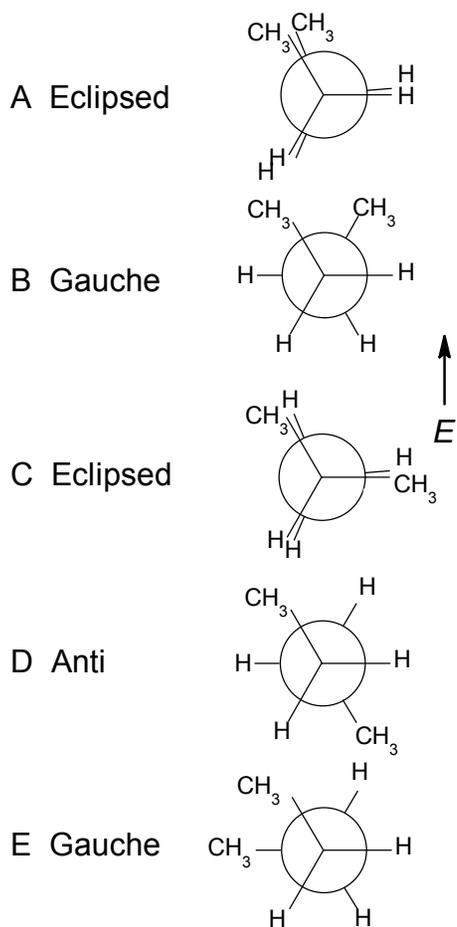
Termination:
Gradient,
0.05
kcal/mol

Initial
Optimization
: Simplex

Bond Stretching Energy : 262.467
Angle Bending Energy : 125.020
Torsional Energy : 8.902
Str-Bend Energy : 2.266
Out of Plane Bending Energy : 0.242
1-4 van der Waals Energy : 13.430
van der Waals Energy : 2.285
1-4 Electrostatic Energy : 8.111
Electrostatic Energy : -19.471

Bond Stretching Energy : 1.554
Angle Bending Energy : 4.657
Torsional Energy : 1.605
Str-Bend Energy : 0.282
Out of Plane Bending Energy : 0.000
1-4 van der Waals Energy : 16.267
van der Waals Energy : 0.506
1-4 Electrostatic Energy : 5.516
Electrostatic Energy : -18.934

Conformational Analysis



Molecules are not rigid

They exist in conformers

For example

70% anti-trans

30% gauche form

Biological activity of a drug molecule is supposed to depend on one unique conformation

Made using ACDLabs
ChemSketch

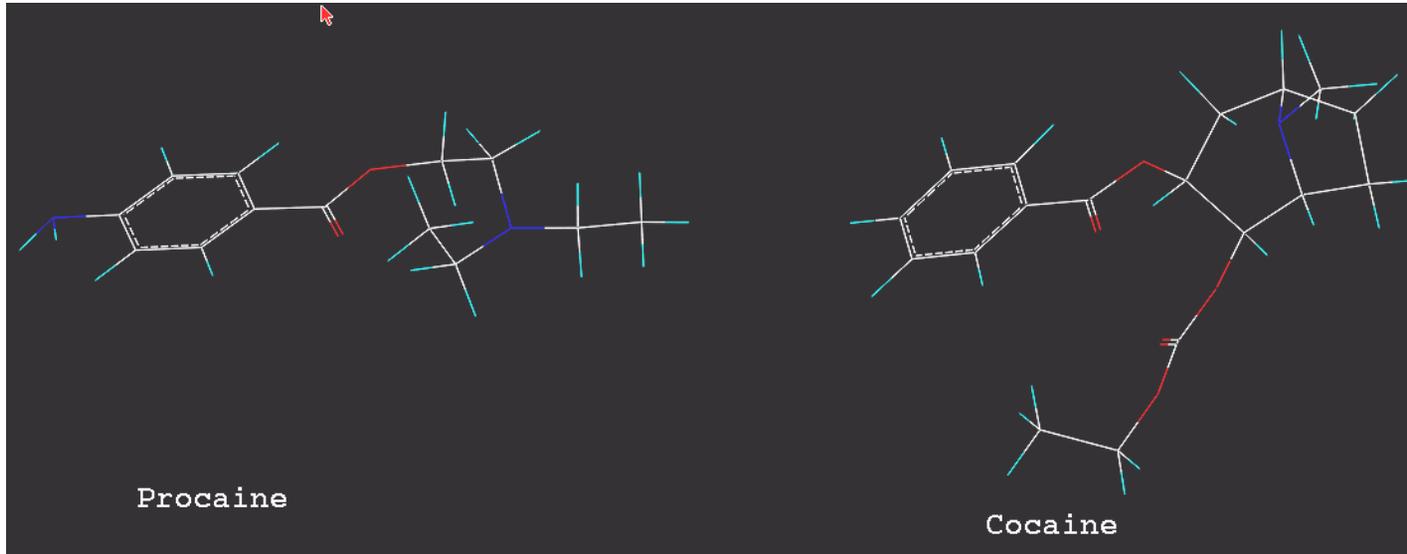
Identifying Active Conformation?

- X-ray structure
 - Crystal structure of target protein with the ligand (drug)
 - Not all proteins can be crystallized (eg. membrane proteins)
- If active compound is a rigid molecule (not many conformations). Take the trial compounds and identify conformations that matches the template
 - MD to identify the conformation

3D Structure Comparison

- Protein Structure Overlays
 - Homology Modeling, Evolutionary relationship, 3D-Folds
- Overlay Small Molecule– Why?
 - Only one conformation results in binding with the receptor. Identifying that active conformation is thus important.
 - Overlay can tell us how two molecules are similar

Pharmacopore



Cocaine and Procaine have anesthetic property.

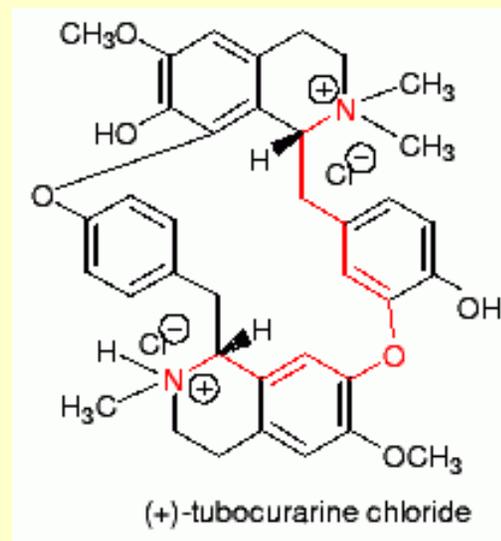
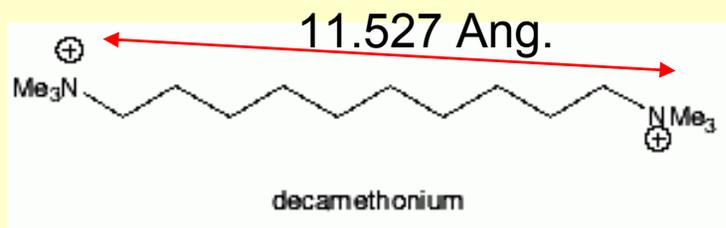
QSAR studies indicate that the pharmacoporic (binding site) is related to the presence of ester, amine and aromatic groups.

Pharmacopore here indicates not only the presence of same functional groups but also their presence in the same relative position

An Introduction to Medicinal
Chemistry, Graham Patrick
2002

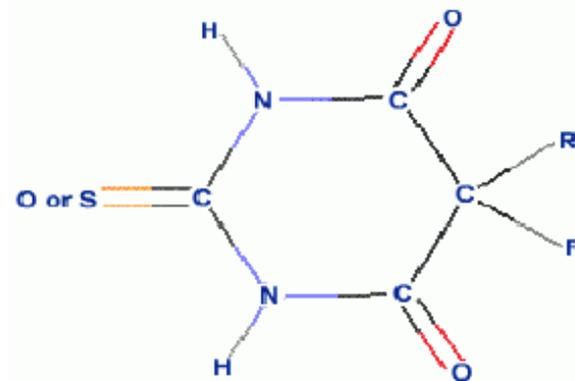
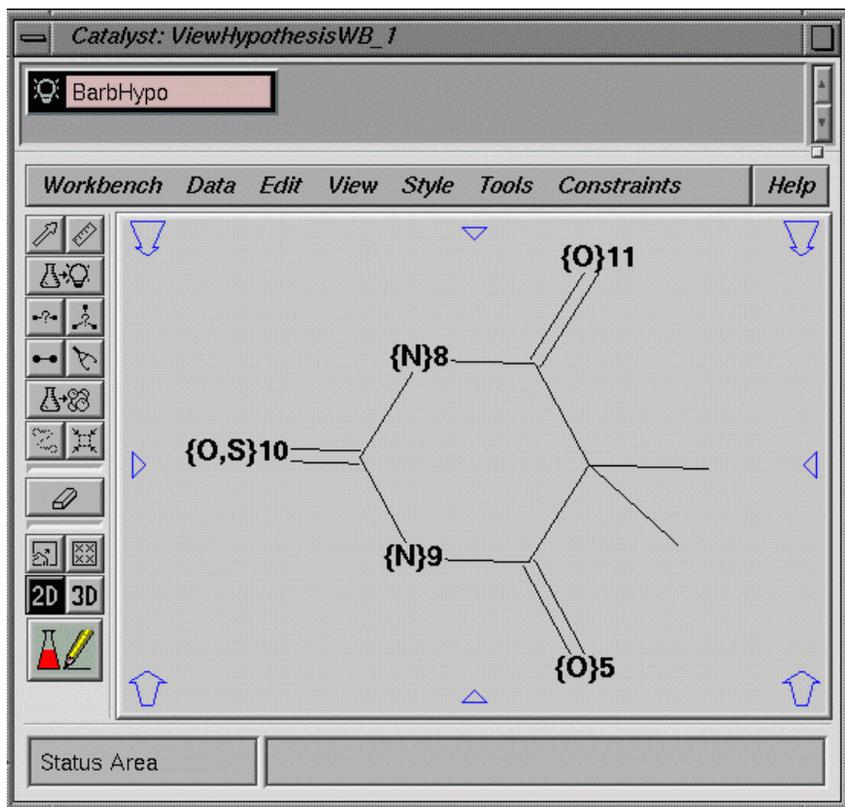
Pharmacophore

Neuromuscular blocking agent-
Pharmacophore: 2 Quaternary N atoms



An Introduction to Medicinal
Chemistry, Graham Patrick
2002

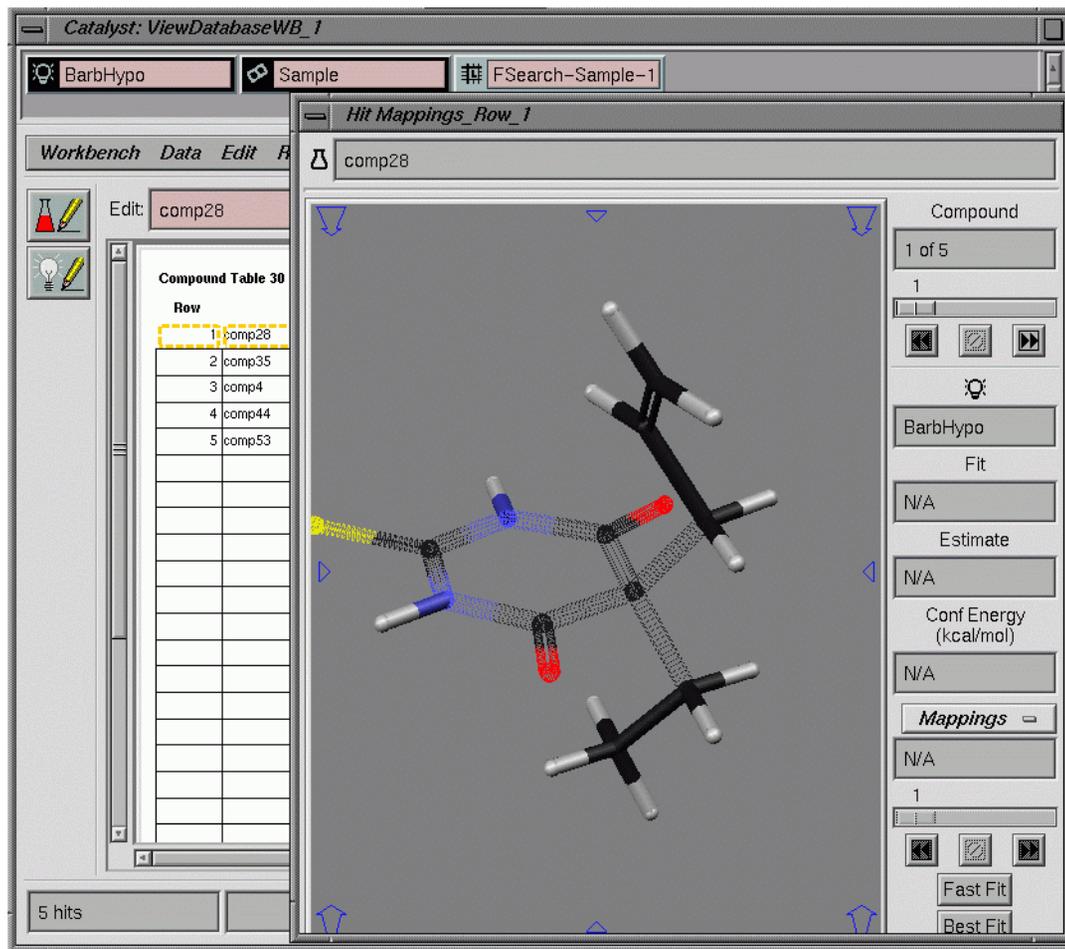
Substructure Search



Substructure for generic
barbiturate

Catalyst 4.9 Accelrys Inc.

Substructure Search



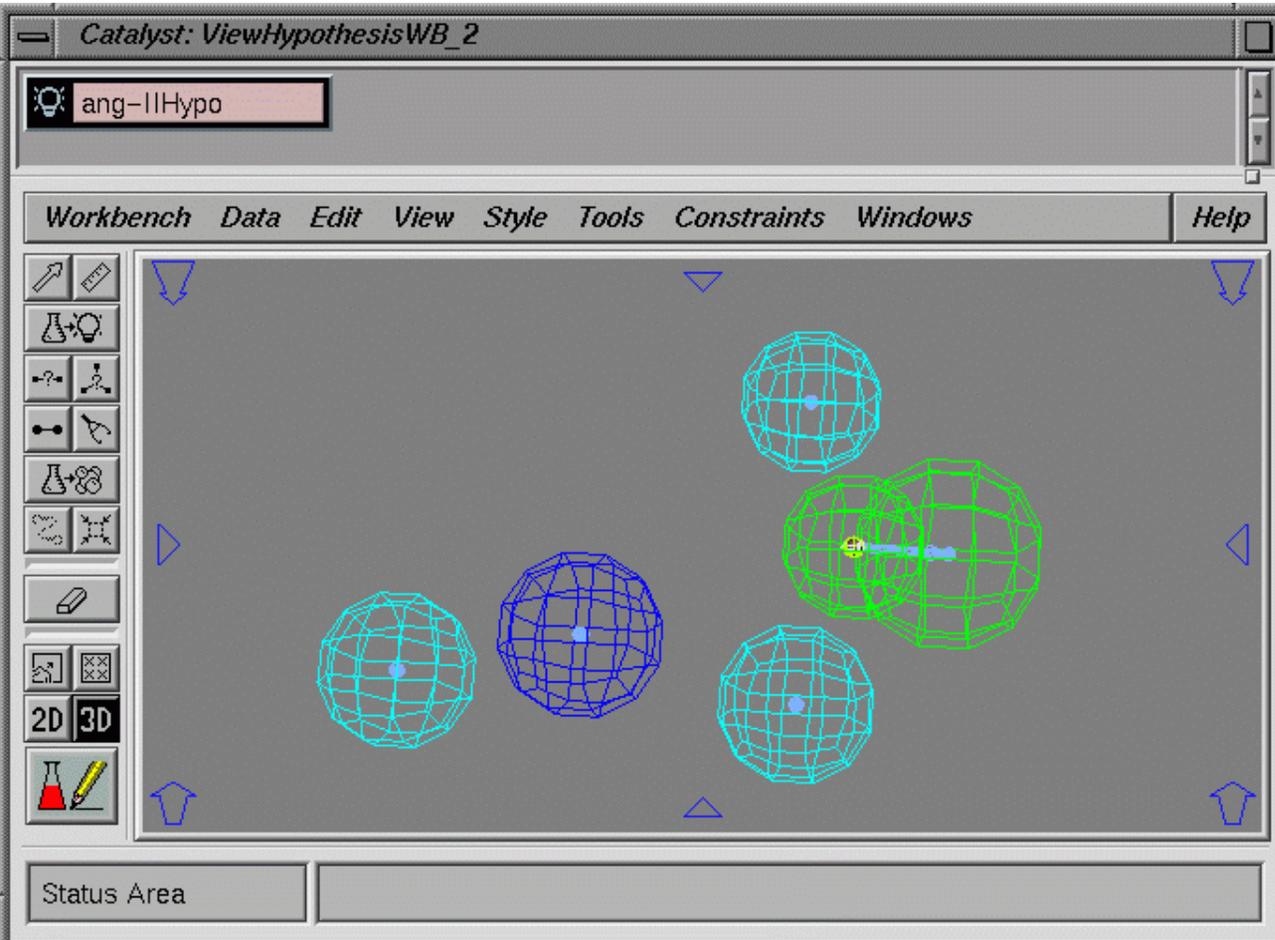
Hits from the **Fast Flexible Search Database/Spreadsheet Search** on a sample database

Options to save the hits

Fit the hits to the hypothesis

Figure shows Catalyst 4.9 interface from Accelrys Inc.

3D-Search



Cyan Hydrophobes
Blue Neg Ionizable
Green HB acceptor

Once you have identified a set of compounds that exhibited activity for the same assay. Use them to generate a 3D hypothesis

Figure shows Catalyst 4.9 interface from Accelrys Inc.

Molecular Dynamics (MD)

- Time dependent behavior of the molecular system
 - Local vibrations, conformational change of proteins and nucleic acids
- MD is based on classical Newton's motion
 - Equation of motion: $F_i(t) = m_i a_i(t)$
 - Gradient of potential energy is used to calculate the forces
- Time-Step
 - Δt
- Several algorithms are available
 - Verlet, Velocity verlet etc.

Gromacs, Amber, Charmm, VMD, NEMD

MD Overview

System

Interaction Potential

$$U_{ij} = 4\epsilon_{ij} \left[\sigma_{ij}/r_{ij}^{12} - \sigma_{ij}/r_{ij}^6 \right]$$

Newton's Equation

$$d^2x_i/dt^2 = F_{xi}/m_i$$

Differential equations are solved using finite-difference methods

At time t : X , V and other dynamic information (known)

Predict at Δt , X , V , etc at reasonable accuracy

Time Step

Δt



Analysis: Correlation Functions etc.

MD

- A typical MD run consist of the following steps
 - Set Initial configuration/Velocity; Heating, Equilibration, Production, Saving configurations
- Applications: Dynamical Properties, MD can take information from NMR to perform a restrained MD.

Quantum Mechanics (QM)

- Need for QM
 - MM and MD do not consider electrons explicitly (Born-Oppenheimer approximation)
 - When a drug molecule interact with a receptor. Primary interactions occur between the electron clouds. ELECTRONIC influence cannot be ignored always.
 - MM and MD cannot answer questions related to
 - Bond-forming or bond-breaking
 - Molecules not in ground state

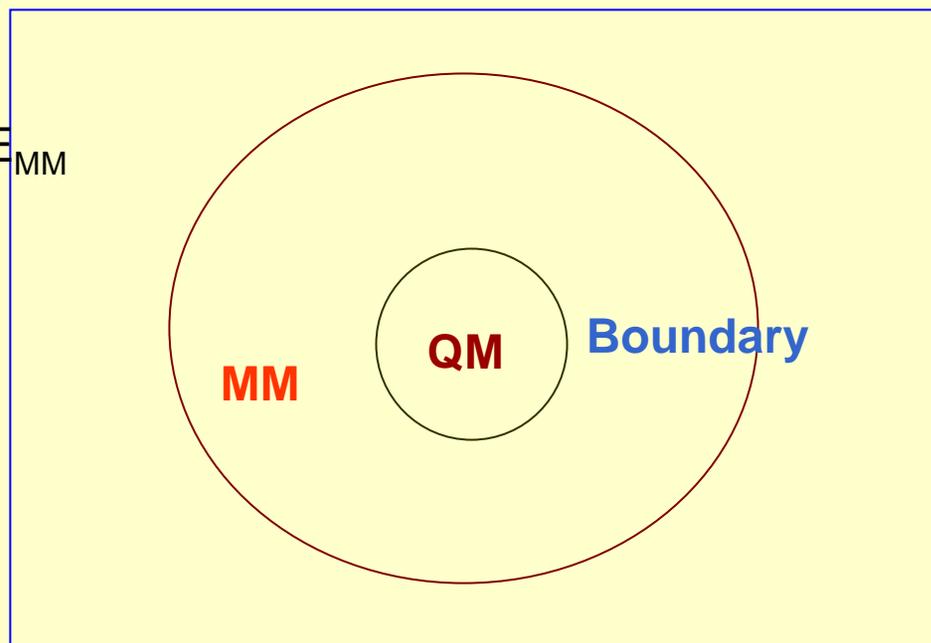
QM

- Basics: $H\Psi = E\Psi$ Shrodinger's Equation E =Energy, Ψ = Wave Function.
 - Solve S.E to get the Energy and Wave Function, which inturn can be used to extract electronic properties (electron density etc.)
 - *ab-initio (from first principles)*, semi-empirical (approximations involved, for ex., only valence electrons are considered (AM1, PM3 etc.)
- QM can be used in conformational search and energy minimization
- Flavors: MOPAC, GAMESS etc.
- Applications: Minimization for small molecules, for conjugated systems, Descriptors for QSAR, Partial charges, transition state geometries & energies

Advanced Techniques: QM/MM

- MC (Monte-Carlo), Brownian Dynamics, QM/MM

$$E = E_{\text{QM,elec}} + E_{\text{QM,vdW}} + E_{\text{MM}}$$



CHARMM has MM/QM module

Advances in Molecular Modeling/MD

Year	System	Total Time	Computer
1983	DNA, Vacuum 12 and 24 bp (754/1530 atoms)	0.09 ns	Several weeks each on Vax780
2002	Channel Protein in lipid membrane (106,189 atoms,PME)	5.00 ns	30 hrs on a 500 proc. LeMieux terascale System 50 days, 23 proc Linux (Athlon)

1 nanosecond = 1×10^{-9} seconds

Molecular Modeling and Simulation
Tamar Schlick

Applications of Molecular Modeling

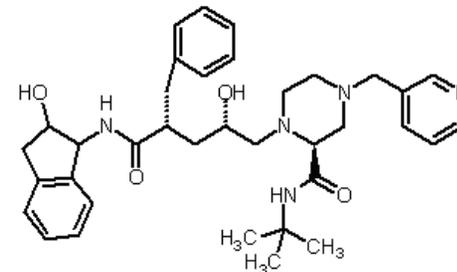
- Protein-Ligand Docking:
 - If you are interested in finding out how small molecules (drug molecule) interact and bind with a receptor of known 3-D structure
 - [AutoDock](#), [Dock](#), [Flexi-dock](#) etc.
- Protein-Protein Docking
 - Rigid body docking
 - 3D-Dock

Applications: Drug Design

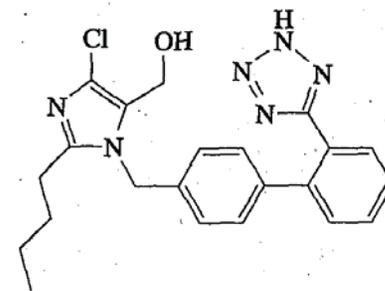
“A model must be wrong, in some respects, otherwise it would be the thing itself. The trick is to see where it is right”

Henry A. Bent

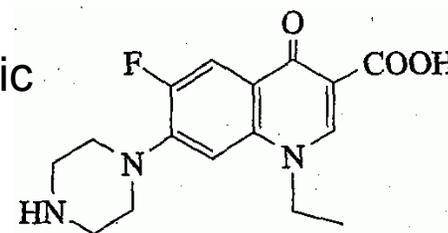
Indinavir: Antiviral Drug/HIV/MM/MD/Crystallography



Losartan: Antihypertensive agent/Structure Activity Study



Norfloxacin: Antibiotic /QSAR



Hands-on Exercise

- Instructions in the web-link
 - <http://nciiris.ncifcrf.gov/~ravichas/MM/MM.htm>

Selected Reference Books (Molecular Modeling)

- Molecular Modeling: Basic Principles and Applications, H.-D. Holtje, W. Sippl, D. Rognan and C. Folkers, second Edition, Wiley-VCH (2003)
- Molecular Modeling and Simulation, T. Schlick (2002)
- Molecular Modeling: Principles and Applications A.R. Leach (2001)
- Computer Simulation of liquids, M.P. Allen and D.J. Tildesley (1989)
- Discovering Genomics, Proteomics & Bioinformatics, A. M. Campbell and L. J. Heyer (2003)
- Bioinformatics: A practical Guide to the analysis of Genes and Proteins, Edited by A.D. Baxevanis and B.F.F. Quellette (2001)

Selected Reference Books (Molecular Modeling)

- Developing Bioinformatics Computer Skills, C.Gibas and P. Jambeck (2001)
- Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins, Andreas D. Baxevanis, B.F. Oullette (2001)
- Introduction to Bioinformatics, Arthur M.Lesk (2002)
- Bioinformatics Basics, H. H. Rashidi and L.K. Buehler (2000)
- Introduction to Bioinformatics: Atwood and Parry-Smith (1999)
- *Chemoinformatics: A Textbook*, J. Gasteiger, T. Engel